

Code No: 155AT

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech III Year I Semester Examinations, March - 2024

DATA ANALYTICS

(Common to CSE, IT)

Time: 3 Hours

Max. Marks: 75

Note: i) Question paper consists of Part A, Part B.

ii) Part A is compulsory, which carries 25 marks. In Part A, Answer all questions.

iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

PART – A**(25 Marks)**

- 1.a) What do you mean by missing values? [2]
- b) Which source of data is more reliable? Why? [3]
- c) What is the objective of reporting? [2]
- d) What is the difference between nominal and ordinal? [3]
- e) What is residual? [2]
- f) How to compute cost function in linear regression? [3]
- g) What is supervised learning? [2]
- h) What is the use of gini index? [3]
- i) What is data visualization? [2]
- j) What are the benefits of using stick figures? [3]

PART – B**(50 Marks)**

2. Explain the different steps in data processing. [10]

OR

3. How to handle noise and outliers in data management? Explain with examples. [10]

4. Provide an overview of data modeling techniques and their application in business. [10]

OR

5. Describe different tools and environments for performing data analytics. [10]

6. Demonstrate ordinary least squares method for linear regression. [10]

OR

7. Make a comparison of linear regression with logistic regression. [10]

8. What is a decision tree? Explain CHAID algorithm for building decision trees. [10]

OR

- 9.a) Cleaning in ETL performs basic data unification. List the rules associated with this process.

- b) Discuss the measures for time series data forecast accuracy. [5+5]

10. Explain geometric projection visualization techniques. [10]

OR

11. Discuss the challenges involved in visualizing complex data and suggest suitable techniques. [10]